What if you could learn Spanish *vocabulario* just by reading enjoyable *párrafos* like this one? People learn new words *en contexto* all the *tiempo*. So *nosotros* have helped you by keeping *mucho* of English context. *Nuestro* model says *que* when *tu* see a *frase*, *tu* *modificas* the embeddings of all new words to improve cloze prediction of *todos* tokens in the sentence. It predicts that *tu* will learn embeddings *para* these Spanish words, in these *contextos*, that *son* similar to *las* embeddings of the original *inglés* words. *Nosotros* constructed this paragraph to *ayudar* you learn as *más* vocabulary as *posible*, using best-first search *para* choose a *subconjunto* de *palabras* *para* *flip*.

**RESEARCH GOAL**

- **Incidental Learning** is vital for native language (L1) acquisition (Krashen, 1989) and L2 acquisition (Huckin & Coady, 1999).
- **Challenge**: For purely L2 incidental learning we (adults) must currently read below our “interest levels”.
- **Our Goal**: Construct macaronic texts that are predicted to enable foreign vocabulary learning, without any human feedback.

**SETUP**

![Image of setup diagram](image_url)

**PROXY STUDENT MODEL**

How **predictable** and **predictive** are novel (L2) words in a sentence?

The embedding of “contexto” should make it more **predictable**.

People learn *nuevos* words *en contexto* all the *tiempo*

and also more **predictive** of the embeddings of all other tokens, e.g.; the embedding of “words”.

We use a **cloze language model** to model predictability and predictivity of tokens initially trained on a L1 corpus.

... and then **incrementally learn embeddings of (L2) words** in a macaronic sentence:

\[
\begin{align*}
L'(\theta) = & \sum \log p(x_t | [h'_t ; h'_p]) \\
= & \sum \log \frac{\text{softmax}(E(F) - h'_t : h'_p)}{\text{softmax}(E(F) - h'_t : h'_p)} \\
= & L(\theta, F) - \lambda |F - F_{\text{prev}}|^2
\end{align*}
\]

**FUTURE WORK**

- Explore other spelling-aware representations (CNNs, RNNs etc.)
- Extend model to allow reordering, partial word and phrasal substitutions.
- Adapt online to student feedback in order to build personalized proxy student models and also model student’s learning and forgetting patterns similar to our prior work in Renduchintala et al. 2017.

**EXPERIMENTS (HUMAN EVALUATIONS)**

- Are macaronic texts useful for human learners of an L2?
- Native English (L1) Mturk “students” were given Spanish/German (L2) infused macaronic texts (short passages from novels) and we measured:
  - a) comprehension of L2 vocabulary in context (participants guessed the meaning of L2 words in macaronic sentence as they read the sentence)
  - b) short-term retention without context (guessed the meanings of isolated L2 words at the end of the passage)

**SEARCH**

We score macaronic configurations by how much they improve F:

\[
MRR(F) = \frac{1}{|F|} \sum_{i=1}^{|F|} \left( \frac{1}{r_{ci} - \text{rank}(F, c, \text{sorted})} \right)
\]

Rank threshold heuristic

Best-first left-to-right search (flipping one token at a time)

*People learn new words in context*

**FUTURE WORK**

- Explore other spelling-aware representations (CNNs, RNNs etc.)
- Extend model to allow reordering, partial word and phrasal substitutions.
- Adapt online to student feedback in order to build personalized proxy student models and also model student’s learning and forgetting patterns similar to our prior work in Renduchintala et al. 2017.